ChinaXiv合作期刊

A new method of searching for concealed Au deposits by using the spectrum of arid desert plant species

CUI Shichao[1,2,3,4], ZHOU Kefa[1,2,3,4*], ZHANG Guanbin[5*], DING Rufu[6], WANG Jinlin[1,2,3,4], CHENG Yinyi[1,2,3,4], JIANG Guo[1,2,3,4]

[1] State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, China;
[2] Xinjiang Key Laboratory of Mineral Resources and Digital Geology, Urumqi 830011, China;
[3] Xinjiang Research Centre for Mineral Resources, Chinese Academy of Sciences, Urumqi 830011, China;
[4] University of Chinese Academy of Sciences, Beijing 100049, China;
[5] Xinjiang Academy of Science and Technology for Development, Urumqi 830011, China;
[6] China Non-Ferrous Metals Resources Geological Survey, Beijing 100012, China

Abstract: With the increase of exploration depth, it is more and more difficult to find Au deposits. Due to the limitation of time and cost, traditional geological exploration methods are becoming increasingly difficult to be effectively applied. Thus, new methods and ideas are urgently needed. This study assessed the feasibility and effectiveness of using hyperspectral technology to prospect for hidden Au deposits. For this purpose, 48 plant (*Seriphidium terrae-albae*) and soil (aeolian gravel desert soil) samples were first collected along a sampling line that traverses an Au mineralization alteration zone (Aketasi mining region in an arid region of China) and were used to obtain soil Au contents by a chemical analysis method and the reflectance spectra of plants obtained with an Analytical Spectral Device (ASD) FieldSpec3 spectrometer. Then, the corresponding relationship between the soil Au content anomaly and concealed Au deposits was investigated. Additionally, the characteristic bands were selected from plant spectra using four different methods, namely, genetic algorithm (GA), stepwise regression analysis (STE), competitive adaptive reweighted sampling (CARS), and correlation coefficient method (CC), and were then input into the partial least squares (PLS) method to construct a model for estimating the soil Au content. Finally, the quantitative relationship between the soil Au content and the 15 different plant transformation spectra was established using the PLS method. The results were compared with those of a model based on the full spectrum. The results obtained in this study indicate that the location of concealed Au deposits can be predicted based on soil geochemical anomaly information, and it is feasible and effective to use the full plant spectrum and PLS method to estimate the Au content in the soil. The cross-validated coefficient of determination (R^2) and the ratio of the performance to deviation (RPD) between the predicted value and the measured value reached the maximum of 0.8218 and 2.37, respectively, with a minimum value of 6.56 μg/kg for the root-mean-squared error (RMSE) in the full spectrum model. However, in the process of modeling, it is crucial to select the appropriate transformation spectrum as the input parameter for the PLS method. Compared with the GA, STE, and CC methods, CARS was the superior characteristic band screening method based on the accuracy and complexity of the model. When modeling with characteristic bands, the highest accuracy, R^2 of 0.8016, RMSE of 7.07 μg/kg, and RPD of 2.20 were obtained when 56 characteristic bands were selected from the transformed spectra $(1/\ln R)'$ (where it represents the first derivative of the reciprocal of the logarithmic spectrum) of sampled plants using the CARS method and were input into the PLS method to construct an inversion model of the Au content in the soil. Thus, characteristic bands can replace the full spectrum when constructing a model for estimating the soil Au content. Finally, this study proposes a method of using plant spectra to find concealed Au deposits, which may have promising application prospects because of its simplicity and rapidity.

*Corresponding authors: ZHOU Kefa (E-mail: zhoukf@ms.xjb.ac.cn); ZHANG Guanbin (E-mail: zhanggb2021@163.com)
Received 2020-09-28; revised 2021-04-11; accepted 2021-04-26
© Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Science Press and Springer-Verlag GmbH Germany, part of Springer Nature 2021

chinaXiv:202112.00013v1

ChinaXiv合作期刊

Keywords: concealed Au deposits; reflectance spectroscopy; soil Au content; characteristic band; soil geochemical prospecting; competitive adaptive reweighted sampling; *Seriphidium terrae-albae*

chinaXiv:202112.00013v1

1 Introduction

Soil geochemical prospecting involves studying the regularity of the dispersion and concentration of metallic elements in the soil and the corresponding relationship with ore bodies in the bedrock. This approach is based on systematic measurements of the distribution of metallic elements in the soil and then prospecting for hidden ore bodies by identifying chemical anomalies and interpreting evaluation anomalies (Yang et al., 2018). Soil geochemical methods have been widely used in geological prospecting because they offer advantages such as strong reliability, low cost, simple implementation, and short operation period (Arias, 1996; Smee, 1998; Smith et al., 2011; Wang et al., 2013; Timofeev et al., 2016). However, in traditional soil geochemical prospecting, the metal content in the collected soil samples is mainly measured by laboratory chemical methods (Von Steiger et al., 1996; Kemper and Sommer, 2002); although the accuracy is high, this process is time consuming and laborious, and the heavy metal content in the soil is only determined in a small area (Yousefi et al., 2018; Han et al., 2020). Additionally, it is difficult to obtain continuously distributed information related to heavy metals at large scales, resulting in high levels of uncertainty regarding the exploration of concealed deposits (Liu et al., 2019).

The emergence of hyperspectral remote sensing technology has provided new methods to macroscopically and rapidly acquire soil heavy metal information over large areas (Sun and Zhang, 2017; Shi et al., 2018). Hyperspectral remote sensing provides multiband, high-spectral-resolution, and continuous results. Thus, continuous soil spectra in the visible, near-infrared, and mid-infrared ranges can be obtained (Cheng et al., 2019). Additionally, heavy metal elements are usually adsorbed or hosted in clay minerals, iron oxides, and organic matter, and exhibit relatively significant spectral characteristics in the soil spectrum; therefore, hyperspectral techniques can potentially be used to remotely retrieve information related to metal elements in the soil (Vega et al., 2006; Sun et al., 2018). Scholars have increasingly used mathematical regression methods to establish the quantitative relationships between heavy metals and spectral characteristic parameters (Kooistra et al., 2001; Kemper and Sommer, 2002; Moros et al., 2009; Chakraborty et al., 2015; Sawut et al., 2018; Yousefi et al., 2018). Moreover, many researchers have predicted the spatially continuous distribution of heavy metals in the soil (Tan et al., 2021; Wang et al., 2021).

Although the above studies have verified that hyperspectral techniques have considerable potential for predicting the metal content in the soil, it is extremely difficult to use these methods to retrieve the metal content in the soil because of interference from plants in plant-covered areas. However, many studies have shown that several plants can absorb metal elements from the soil through their roots during growth (Bandaru et al., 2016; Cui et al., 2021). When heavy metal elements excessively accumulate in plants, the chlorophyll content, cell structure, and water vapor content of the plants change, resulting in changes to their reflectance spectra (Hoque and Huntzler, 1992; Dunagan et al., 2007; Ren et al., 2008), which can provide a theoretical basis for the indirect inversion of metal element contents in the soil. Some scholars have successfully estimated the heavy metal content in the soil beneath plants by using plant spectra. For instance, Kooistra et al. (2003) found that a vegetation index (Modified Soil-Adjusted Vegetation Index 2 Model (MSAVI2)) constructed based on the reflectance spectra of perennial ryegrass (*Lolium perenne*) could be used to characterize the level of Zn pollution in the soil. Shi et al. (2016a) studied rice and found that a three-band vegetation index displayed greater potential than a two-band vegetation index for estimating the metal content in the soil because of its ability to use more spectral information. Among all the various indices, the three-band vegetation index $((R_{716}-R_{568})/(R_{552}-R_{568})$ (where R_{716}, R_{568}, and R_{552} represent the reflectivity at 716, 568, and 552 nm wavelengths, respectively) is considered the optimal index for estimating soil metal contents. The above studies have shown that plant spectra and indices

ChinaXiv合作期刊

constructed based on spectra can be used to effectively monitor the heavy metal content in the soil in plant-covered areas.

Arid and semi-arid regions in China is rich in mineral resources, but it is difficult to find valuable mineral deposits there due to the presence of thick Quaternary formations. Thus, new prospecting methods and ideas are urgently needed. A small half-shrub species called *Seriphidium terrae-albae* is widely distributed in the mining area, which encompasses the Aketas gold-copper mine, the Kalatongke copper-nickel mine, the Xilekudute copper-molybdenum mine, and the Ketebieteti copper-nickel mine. Scholars have verified that the plant is able to absorb Au from the soil without hindrance and displays a geochemical anomaly (Song et al., 2016; Song et al., 2017). Such an anomaly may provide a reference for the exploration of hidden Au deposits. Although extensive past research has demonstrated that plant spectra can be used to accurately estimate the contents of metals in the soil (Kooistra et al., 2003; Shi et al., 2016a, b), plant species and soil type are also two key factors related to the plant response to metals (Horler et al., 1980). Therefore, it is necessary to further verify whether the spectrum of *Seriphidium terrae-albae* can be used to estimate the Au content in arid desert soils. To the best of our knowledge, this topic has not been addressed.

Therefore, the first goal of this study is to determine whether abnormalities in the soil Au content can be used as a prospecting indicator for concealed Au deposits; if so, can they indicate the location of deep hidden ore deposits? The second purpose of this study is to determine whether the spectrum of a plant species (*Seriphidium terrae-albae*) widely distributed in the arid desert area can be used to estimate the Au contents in the soil with high accuracy. If these two tasks are proven feasible, it would suggest that plant spectra can be used to quickly identify the hidden Au deposits. This finding would be of considerable significance for expanding the prospecting space, improving the prospecting efficiency, and accelerating the pace of mineral exploration.

2 Methods and materials

2.1 Study area and experimental design

The study area is located in the Aketasi mining region in arid areas of China, where the main mineral types are Au and Cu. The main soil type in the study area is aeolian gravel desert soil (Song et al., 2017). Traditional geological prospecting methods cannot be effectively implemented in this area because quaternary deposits cover most of the area. By contrast, soil geochemical prospecting has notable potential in this area due to its deep penetration ability.

The main plant species distributed in the study area is *Seriphidium terrae-albae* (as shown in Fig. 1), which is a small half-shrub with small leaves and a relatively developed root system; it can grow in extremely dry and harsh environments because of its high resistance to drought, high temperature, sand damage, and poor soil. It is a small half-shrub that usually germinates in late March and grows rapidly, reaching a peak in mid-May. After a brief summer dormancy, it germinates in early August, blooms in mid-August, and seeds in September before ripening in late October. The plant species enters a drought stage in November and effectively makes it through the winter.

Fig. 1 Natural landscape of the study area

chinaXiv:202112.00013v1

ChinaXiv合作期刊

To determine whether the Au content in the soil can be estimated using *Seriphidium terrae-albae* plant spectra, we designed a 1-km-long sampling line with 48 sampling points at 20 m intervals (as shown in Fig. 2) in accordance with the locations of a known hidden Au deposit. The sampling line crossed the deposit in the north-south direction and crossed the upper part of the deposit and the background area. The geographic locations of the 48 sampling points were recorded using a handheld global positioning system (GPS). The purposes of this design are to ensure that the range of Au contents in the collected soil samples is wide and to obtain representative samples. Next, the plant and soil samples were collected along the sampling line. Healthy plants (*Seriphidium terrae-albae*) growing around the sampling points were sampled, and the weight of each sample was measured as 200 g. The location of each soil sample was the same as that of a corresponding plant, and each soil sample was also weighed approximately 200 g.

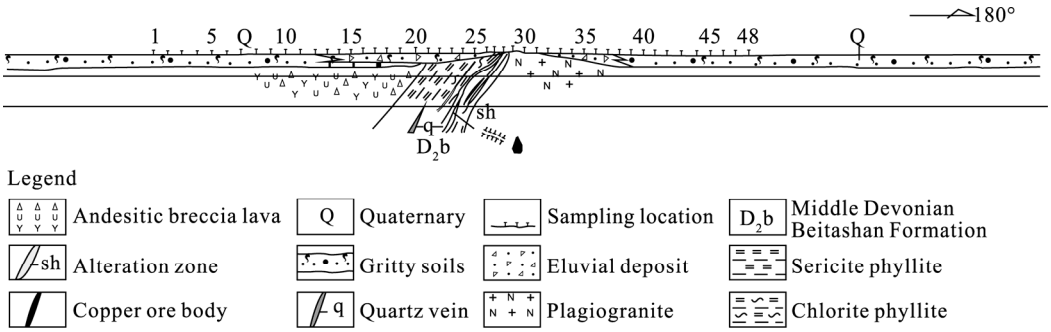

Fig. 2 Demonstration of the locations of sampling points. Note that the numbers represent the sampling points.

2.2 Measurement of the Au content in the soil samples

The collected soil samples were packaged in fresh-keeping bags and taken to the laboratory. They were dried, ground, and stored in sealed containers to measure the heavy metal content. Before element analysis, the soil samples were dissolved using the nitric acid-hydrogen peroxide digestion method, and then the Au content in the soil samples was measured with a ZEEnit650P atomic absorption spectrometer produced by Analytik Jena AG Company in Jena, Germany. The detection limit of the ZEEnit650P atomic absorption spectrometer is 0.2 ng/g. The statistical results for the Au content in the 48 soil samples collected are shown in Table 1.

Table 1 Statistics of the Au content in the 48 collected soil samples

Element type	Minimum value (μg/kg)	Maximum value (μg/kg)	Average value (μg/kg)	Standard deviation (μg/kg)	Coefficient of variation
Au	0.3437	77.5711	6.0543	15.7080	2.5945

2.3 Collection of plant spectra and preprocessing

The canopy plant reflectance spectra were measured using an Analytical Spectral Device (ASD) Fieldspec3 portable spectrometer (Analytical Spectral Devices, Boulder, CO, USA) in clear and windless weather. To reduce the influence of frequent changes in the elevation angle of the sun on the spectral measurements, we selected the measurement period from 12:00 to 14:00 Beijing time on 20 July 2016. During the measurement process, the spectrometer probe was placed vertically downward with a 25° field of view, and the distance between the probe and the plant canopy top was 0.2 m. Five reflectance curves were first collected from plant samples at each sampling point; then, spectral curves that highly varied from the others were removed. Finally, the average value of the remaining spectral curves was taken as the final reflection spectrum of each plant sample at each sampling point.

Reflectance spectra in the range of less than 400 nm and greater than 2400 nm were removed to reduce noise (Sun and Zhang, 2017). Reflectance spectra in the ranges of 1300–1400 and 1800–2000 nm were also removed to eliminate the influence of atmospheric water vapor (Wang et al., 2015). Finally, each sample included 1700 bands.

chinaXiv:202112.00013v1

ChinaXiv合作期刊

The adjacent bands of hyperspectral data are highly correlated, leading to information redundancy. Therefore, in this study, the original spectra were resampled by calculating the average value of five adjacent bands. The reflectance spectrum of each plant sample was reduced to 340 bands after resampling. The reflectance spectra of the plant samples collected after pretreatment are shown in Figure 3.

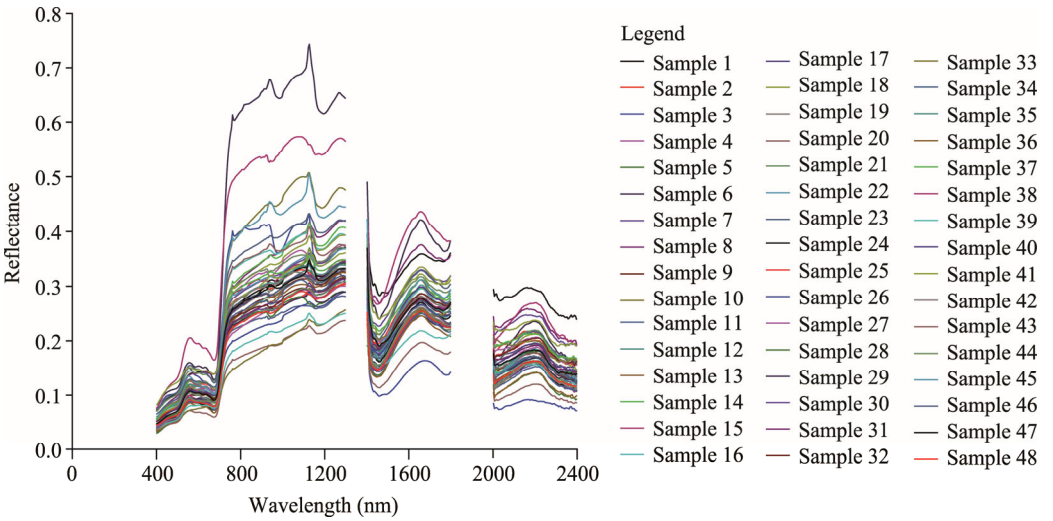

Fig. 3 Reflectance spectra of the collected 48 plant samples

2.4 Partial least squares (PLS) method

The PLS method is a multiple linear regression approach that combines the advantages of principal component analysis, typical correlation analysis, and multiple linear regression analysis (Han et al., 2020). This method is especially suitable for modeling when there is multicollinearity between independent variables and the number of samples is less than the number of variables (Huang et al., 2004). The PLS method can simplify the structure of hyperspectral data by effectively reducing the information redundancy and overcoming multiple correlation issues among bands (Jin and Wang, 2019). Additionally, a hyperspectral inversion model established based on the PLS method can maximize the use of spectral information through principal component analysis to improve the modeling accuracy and estimates; therefore, the PLS method was chosen in this study.

During the modeling process using the PLS method, the number of optimal principal components was determined using the coefficient of determination (R^2) between the predicted value obtained through leave-one-out cross-validation method and the measured value. The number of principal components corresponding to the maximum R^2 value was selected as the optimal number of principal components.

The leave-one-out cross-validated R^2, the ratio of performance to deviation (RPD), and the root-mean-squared error (RMSE) were computed for model evaluation. In general, a good model should have high R^2, high RPD, and low RMSE values.

2.5 Spectral transformations

Various spectral transformations, the original spectrum (R), logarithmic spectrum (ln$R$), reciprocal spectrum (1/R), square root of the original spectrum ($\sqrt{R}$), reciprocal of the logarithmic spectrum (1/ln$R$), and the corresponding first and second derivatives, were input into the PLS method as independent variables to construct a quantitative model for estimating the Au content in the soil. The sequence numbers of the 15 transformed spectra and the corresponding calculation methods are as follows: ST_1, original reflectance spectrum (R); ST_2, the first derivative of the original reflectance spectrum (R'); ST_3, the second derivative of the original reflectance spectrum (R''); ST_4, logarithm of the original spectrum (ln$R$); ST_5, the first derivative of the logarithm of the original spectrum (ln$R$)'; ST_6, the second derivative of the logarithm of the original spectrum (ln$R$)''; ST_7, reciprocal of the

chinaXiv:202112.00013v1

ChinaXiv合作期刊

original spectrum ($1/R$); ST_8, the first derivative of the reciprocal of the original spectrum ($1/R$)'; ST_9, the second derivative of the reciprocal of the original spectrum ($1/R$)''; ST_{10}, square root of original spectrum ($\sqrt{R}$); ST_{11}, the first derivative of the square root of the original spectrum ($\sqrt{R}$)'; ST_{12}, the second derivative of the square root of the original spectrum ($\sqrt{R}$)''; ST_{13}, reciprocal of logarithmic spectrum ($1/\ln R$); ST_{14}, the first derivative of reciprocal of logarithmic spectrum ($1/\ln R$)'; and ST_{15}, the second derivative of reciprocal of logarithmic spectrum ($1/\ln R$)''.

Derivative spectroscopy has been increasingly used in spectroscopic analysis, and this study is no exception. Notably, in the derivatives of plant spectra, the influence of the soil background was effectively removed, and derivatives can be used to accurately extract physiological information from crop canopy spectra (Shi et al., 2014; Cui et al., 2018). In short, derivative spectroscopy can conveniently determine the characteristic parameters of a spectrum, such as the bending points and wavelength positions associated with the maximum and minimum reflectivity (Cui et al., 2018). Thus, subtle spectral changes in soil or plants caused by metal stress can be identified. We used the difference method to calculate the first and second derivatives because the collected plant spectra were discrete data. The formulas for calculating the first and second derivatives are shown in Equations 1 and 2, respectively.

$$S'(\lambda)=\frac{S(\lambda_{i+1})-S(\lambda_{i-1})}{2\Delta\lambda}, \tag{1}$$

$$S''(\lambda)=\frac{S'(\lambda_{i+1})-S'(\lambda_{i-1})}{2\Delta\lambda}, \tag{2}$$

where $S(\lambda_{i+1})$ represents the value at wavelength λ_{i+1} (nm) under different spectral transformations; $S(\lambda_{i-1})$ represents the value at wavelength λ_{i-1} (nm) under different spectral transformations; $S'(\lambda)$ represents the first derivative at wavelength λ (nm) under different spectral transformations; $S''(\lambda)$ represents the second derivative at wavelength λ (nm) under different spectral transformations; and $\Delta\lambda$ (nm) represents the wavelength interval.

2.6 Feature band selection method and evaluation of the constructed model

The reflectance spectra of plants collected using the ASD spectrometer included a vast number of spectral bands, and different bands may make different contributions in the estimation of the Au content in the soil. The characteristic wavelength variables that significantly contribute to the Au content estimates were selected via various methods to reduce the modeling time and simplify the modeling process. The most important goal was to eliminate irrelevant variables and establish a robust and accurate quantitative model.

In this study, four methods, including the genetic algorithm (GA), stepwise regression analysis (STE), competitive adaptive reweighted sampling (CARS), and correlation coefficient (CC) method, were used to select the characteristic bands from all spectral bands in the range of 400–2400 nm; the result was then input into the PLS method to construct an inversion model of the Au content in the soil. Because the number of samples was small, the R^2 between the predicted Au content obtained using leave-one-out cross-validation method and the measured Au content was used to evaluate the accuracy and stability of the constructed model. The larger the R^2 value, the greater the accuracy and stability of the constructed model.

The GA is a stochastic global search and optimization algorithm based on natural selection and natural genetic mechanisms in the biological world (Hong et al., 2018). This algorithm can simulate natural selection and the phenomena of replication, crossover, and mutation in the genetic process. Through continuous genetic iterations, variables that increase the value of the objective function are retained, comparatively poor variables are eliminated, and globally optimal results are ultimately obtained (Jarvis and Goodacre, 2005). Additionally, a parallel search method that can improve the accuracy and stability of the analysis results is used. Therefore, in this study, the GA and partial least squares (GA-PLS) method were combined to estimate the Au content. First, the Au content in the soil was identified as the optimization objective, and plant spectral data were used as genes. Binary coding was then performed to randomly generate the initial population. The root-mean-squared error of cross-validation (RMSECV) was used as the fitness function, and the GA was used to optimize

chinaXiv:202112.00013v1

ChinaXiv合作期刊

the input characteristic bands. The selected characteristic bands were then entered into the PLS method to establish an estimation model of the Au content. In using the GA method for feature band screening, the population size was set to 30, the maximum number of iterations was 100, the crossover probability was 0.5, and the mutation probability was 0.01.

The STE screening of characteristic bands involved first sequentially adding each spectral band to the statistical model and then performing an *F* test of the statistical model after each spectral band was introduced. The *P*-value of the *F* test was used for spectral band introduction and elimination (Jin and Wang, 2019). In this study, the maximum *P*-value for a spectral band to be included was defined as 0.05, and the minimum *P*-value for a spectral band to be removed was defined as 0.10.

In this study, the CC method was also used to select characteristic bands. First, the coefficient and *P*-value between each band within the range of 400–2400 nm and the Au content in the soil were calculated, and then spectral bands with *P*-values less than 0.01 were screened and removed from the selected characteristic bands.

The CARS algorithm simulates Darwin's "adaptive survival" principle (Li et al., 2009; Hong et al., 2018). In the process of characteristic band selection using this method, each wavelength variable was individually regarded, wavelength variables associated with regression coefficients with large absolute values in the PLS model were retained, and wavelengths associated with regression coefficients with small absolute values were eliminated by using an adaptive resampling weighting method and an exponential decay function. In this way, many subsets of spectral variables can be screened, and the optimal spectral variable subset can be selected on the basis of the RMSECV obtained using cross-validation. The subset corresponding to the minimum RMSECV was chosen as the optimal variable subset. In this study, the Monte Carlo sampling frequency was defined as 50.

3 Results

3.1 Analysis and evaluation of soil geochemical characteristics

From Figure 4 it can be seen that, there was an M-shaped soil Au content anomaly in the area corresponding to sampling points 25–33. Furthermore, Figure 5 shows that the soil Au content in this area was significantly higher than that in other areas ($P<0.001$), which indicates that abnormal soil Au contents may be present. Moreover, concealed gold mineralization zones have been found below these areas, and the drilling data indicate a tenor of ore in the range from 1.5 to 4.0 g/t (Song et al., 2017). The above results imply that the abnormal soil Au contents can potentially indicate the locations of hidden Au deposits. Therefore, using the proposed soil geochemical method for mineral exploration in this area is feasible and effective. However, it is time consuming and laborious to use traditional methods to measure the Au content in the soil. Therefore, in later sections, we used hyperspectral technology to quickly estimate the Au content in the soil.

3.2 Influence of spectral transformation on the estimation of the Au content

Table 2 shows that when using all bands in the range of 400–2400 nm in *Seriphidium terrae-albae* plant spectra and the PLS method to construct the model for estimating the Au content in the soil, there were significant differences in the estimation accuracies of the model for different forms of the transformed spectra. The R^2 between the predicted Au content obtained using the leave-one-out cross-validation method and the measured Au content in the soil varied widely from 0.1785 to 0.8218. The RMSE and RPD values of the estimation models constructed using different forms of the transformed spectra also considerably varied. The lowest RMSE value was 6.56 μg/kg, and the highest was 14.53 μg/kg; additionally, the lowest RPD value was 1.07, and the highest was 2.37. These results indicate that it is crucial to choose the appropriate spectral transformation when establishing a model for estimating the Au content in the soil using plant spectra.

Table 2 also indicates that the Au content estimation model obtained by inputting five spectra ST_1, ST_4, ST_7, ST_{10}, and ST_{13} as independent variables into the PLS method is comparatively inaccurate, with maximum R^2 and RPD values of 0.4451 and 1.31, respectively; the lowest RMSE was as high as 11.87 μg/kg. These results imply that some types of transform spectra are not suitable for

chinaXiv:202112.00013v1

ChinaXiv合作期刊

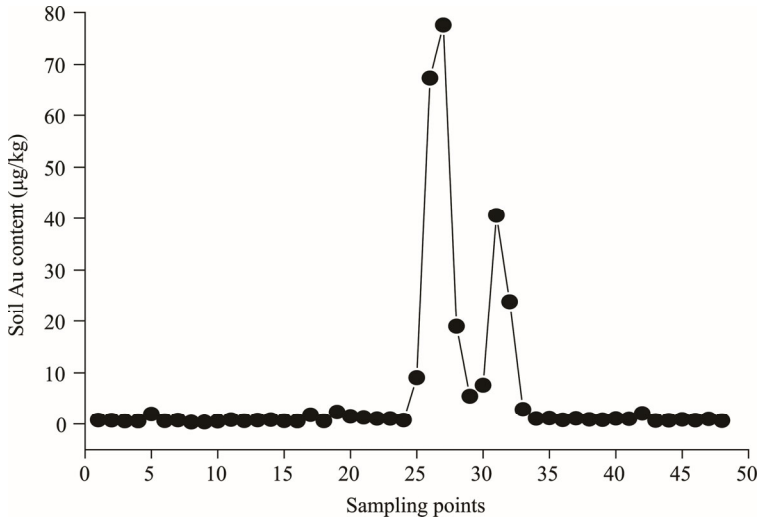

Fig. 4 Distribution of the soil Au content at the 48 sampling points

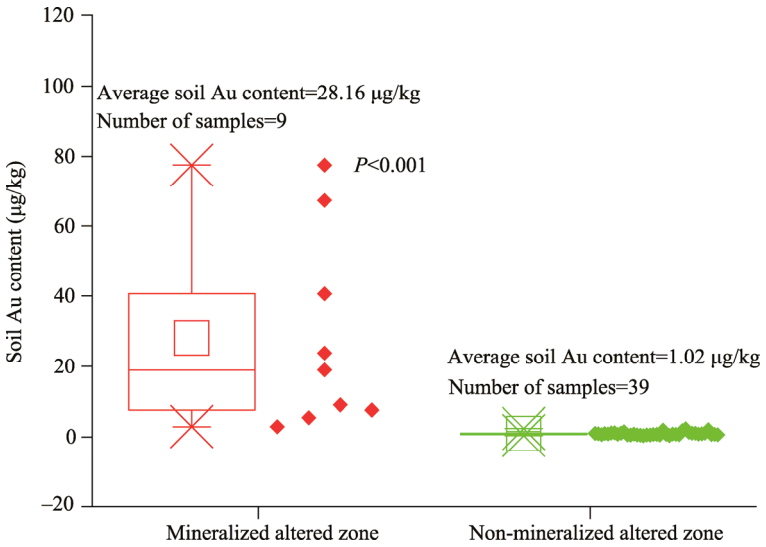

Fig. 5 Number and average Au content of soil samples grown in mineralized and non-mineralized altered zones. Note that boxes represent interquartile ranges (25th to 75th percentiles); thick horizontal bars in each box denotes the median (50th percentile); whiskers (thin horizontal bars) represent the highest and the lowest values, respectively; red and green rectangles denote the average value; and red and green dots represent the Au content of each soil sample in the mineralized altered zone and non-mineralized altered zone, respectively.

estimating the Au content in the soil. By contrast, the accuracy of the estimation model of the soil Au content constructed by inputting the first and second derivatives of these five types of transformed spectra as independent variables into the PLS method was significantly increased. The lowest R^2 and RPD values were 0.2974 and 1.07, respectively, and the highest values even reached 0.8218 and 2.37, respectively. Compared with those of the original spectra, the R^2 and RPD values of the model constructed based on derivative spectra were greatly improved, with increasing ranges of 16.96%–235.32% and –0.926%–80.916%, respectively. Finally, the results indicated that the ST_{14} ((1/ln$R$)') index provides the best spectral transformation. The modeling accuracy in estimating the Au content was highest when using this type of spectral transformation compared to the 14 other types of transformed spectra selected in this study.

Figure 6 shows the scatter diagram of the predicted soil Au content obtained using the leave-one-out cross-validated method versus measured soil Au content. Notably, when the Au content in the

chinaXiv:202112.00013v1

chinaXiv:202112.00013v1

Table 2 Comparisons of the R^2, RMSE, and RPD values obtained using leave-one-out cross-validation method that constructed by 15 different transformed spectra for each of four band selection methods to estimate the soil Au content

Sequence number	Spectral transformation	All bands			GA			STE			CARS			CC		
		R^2	RMSE (μg/kg)	RPD	R^2	RMSE (μg/kg)	RPD	R^2	RMSE (μg/kg)	RPD	R^2	RMSE (μg/kg)	RPD	R^2	RMSE (μg/kg)	RPD
ST_1	R	0.2268	13.82	1.12	0.3594	12.44	1.25	0.1497	14.54	1.07	0.4712	11.44	1.36	0.0988	15.26	1.02
ST_2	R'	0.7605	7.96	1.95	0.8011	7.33	2.12	0.4725	12.98	1.20	0.7068	8.51	1.83	0.5827	10.09	1.54
ST_3	R''	0.6028	10.23	1.52	0.6833	9.16	1.70	0.3496	13.51	1.15	0.5255	10.85	1.43	0.5453	10.66	1.46
ST_4	ln$R$	0.1785	14.43	1.08	0.2128	14.10	1.10	0.0180	17.39	0.89	0.1079	15.22	1.02	0.1504	14.50	1.07
ST_5	(ln$R$)'	0.4260	14.53	1.07	0.4144	14.76	1.05	0.3402	14.40	1.08	0.4350	12.50	1.24	0.4209	19.75	0.79
ST_6	(ln$R$)''	0.4355	11.85	1.31	0.3837	12.68	1.23	0.3277	13.76	1.13	0.2394	14.20	1.09	0.2857	13.18	1.18
ST_7	1/R	0.2259	13.96	1.11	0.3008	13.24	1.17	0.0082	16.57	0.94	0.2824	14.84	1.05	0.1381	14.92	1.04
ST_8	(1/R)'	0.2974	13.05	1.19	0.3330	12.71	1.22	0.3355	12.69	1.22	0.1454	16.02	0.97	0.3014	13.31	1.17
ST_9	(1/R)''	0.3434	12.60	1.23	0.3245	12.83	1.21	0.3115	13.07	1.19	0.3051	12.99	1.20	0.3312	13.14	1.18
ST_{10}	$\sqrt{R}$	0.2056	14.00	1.11	0.2235	13.85	1.12	0.1206	14.90	1.04	0.1828	14.34	1.08	0.0790	15.45	1.01
ST_{11}	($\sqrt{R}$)'	0.5972	10.18	1.53	0.6040	10.05	1.55	0.3934	13.80	1.13	0.4251	12.90	1.20	0.5237	11.08	1.40
ST_{12}	($\sqrt{R}$)''	0.5826	10.26	1.51	0.5237	10.77	1.44	0.3111	12.92	1.20	0.4928	11.10	1.40	0.4789	11.28	1.38
ST_{13}	1/ln$R$	0.4451	11.87	1.31	0.4329	12.33	1.26	0.5312	12.58	1.24	0.5639	10.46	1.49	0.4970	11.50	1.35
ST_{14}	(1/ln$R$)'	0.8218	6.56	2.37	0.7601	7.86	1.98	0.6946	8.87	1.75	0.8016	7.07	2.20	0.7868	8.17	1.90
ST_{15}	(1/ln$R$)''	0.5206	10.84	1.43	0.5459	10.48	1.48	0.3755	18.75	0.83	0.6557	9.20	1.69	0.5318	11.06	1.41

Note: R^2, coefficient of determination; RMSE, root-mean-squared error; RPD, ratio of the performance to deviation; GA, genetic algorithm; STE, stepwise regression analysis; CARS, competitive adaptive reweighted sampling; CC, correlation coefficient; R, original spectrum; ln$R$, logarithmic spectrum; 1/R, reciprocal spectrum; 1/ln$R$, reciprocal of logarithmic spectrum; $\sqrt{R}$, square root of original spectrum; ', first derivative; '', second derivative.

ChinaXiv合作期刊

ChinaXiv合作期刊

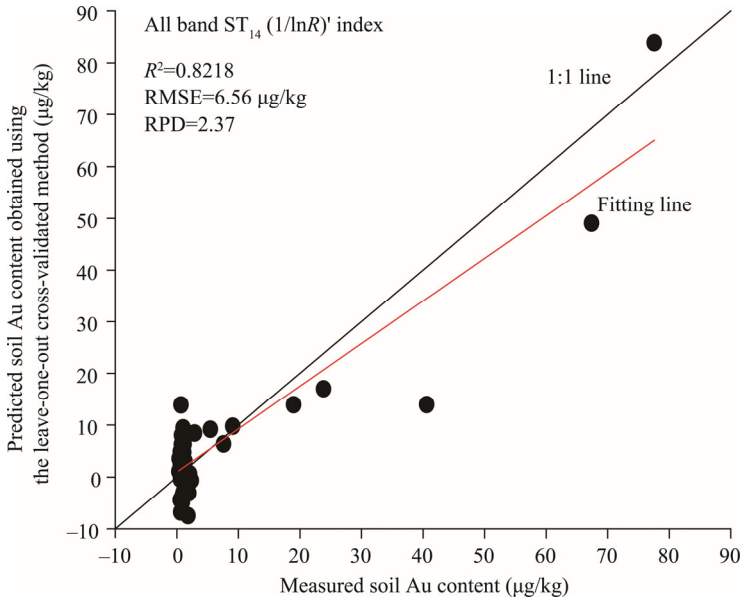

Fig. 6 Scatter diagram of the predicted soil Au content obtained using the leave-one-out cross-validated method versus measured soil Au content. R^2, coefficient of determination; RMSE, root-mean-squared error; RPD, ratio of performance to deviation; 1/ln$R$, reciprocal of logarithmic spectrum; ', first derivative.

soil was greater than 5 μg/kg, the prediction effect was better, and the predicted and measured values were evenly distributed on both sides of the 1:1 line. By contrast, when the Au content was less than 5 μg/kg, the prediction effect was poor, and the predicted Au contents of some samples were negative.

3.3 Influence of the band selection method on the Au content estimations

To study the effects of different band selection methods on the estimation of the Au content in the soil, we first selected the characteristic bands of 15 types of different transformed spectra using the GA, CARS, STE, and CC methods. Then, we input the selected characteristic bands into the PLS model to build an estimation model of the Au content in the soil; constructed 15 estimation models for each band selection method, and calculated the R^2, RMSE, and RPD between the predicted value obtained using the leave-one-out cross-validation method and measured Au content (see Table 2). Finally, we used the average values of the R^2, RMSE, and RPD for the 15 models, which were respectively defined as the M1, M2, and M3 values, as indicators to evaluate the stability and generalization ability of characteristic band selection with different methods. The larger the values of M1 and M3, the smaller the M2 value, and the higher the stability and accuracy of the band selection method.

Figure 7 shows that the M1 and M3 values of the GA method were slightly higher than those of the full spectrum method, and the M2 value of the GA method was slightly lower than that of the full spectrum method. Additionally, the M1 and M3 values of the CARS method were slightly lower than the full spectrum method, and the M2 value of the CARS method was slightly higher than those of the full spectrum method. The M1 and M3 values of the STE and CC methods were considerably lower than those of the full spectrum method, and the M2 value was significantly higher than that of the full spectrum method. If deciding among methods based only on the M1, M2, and M3 values, the GA was the best band selection method. However, in the process of modeling, the number of characteristic bands used was also a key factor that should be considered. Using too many characteristic bands can increase the complexity of the model and reduce its anti-interference capabilities. Therefore, when determining the optimal band selection method, the accuracy and complexity of the constructed model need to be comprehensively considered. Thus, the average numbers of characteristic bands in the 15 estimation models constructed based on four

chinaXiv:202112.00013v1

ChinaXiv合作期刊

band selection methods were also calculated. Figure 7b indicates that the number of characteristic bands selected by the GA method was the largest, reaching as high as 147, or accounting for 43.24% of the 340 total bands; the CC method used the second-largest number of bands. However, the numbers of characteristic bands used in the STE and CARS methods were the lowest, at 14 and 96, respectively, accounting for 4.12% and 15.00% of the 340 total bands. Although the M1 and M3 values of the GA method were slightly higher than those of the CARS method and the M2 value was slightly lower, the number of characteristic bands selected by the GA method was also far greater than that selected by the CARS method. Considering the accuracy and complexity of the model, CARS is the optimal band selection method for the estimation of the Au content in the soil using *Seriphidium terrae-albae* plant spectra.

Figure 7b also suggests that three of the band selection methods, including the GA, CARS, and CC methods, can greatly reduce the number of bands used in the constructed model without greatly reducing the Au content estimation accuracy. The bands used in the models were only in the range of 400–2400 nm, accounting for 15.00%–43.24% of the total number of possible bands. Thus, filtering for noise may greatly reduce the complexity of the model. Nevertheless, the selected characteristic bands can be used to replace all bands in constructing an estimation model of the Au content in the soil.

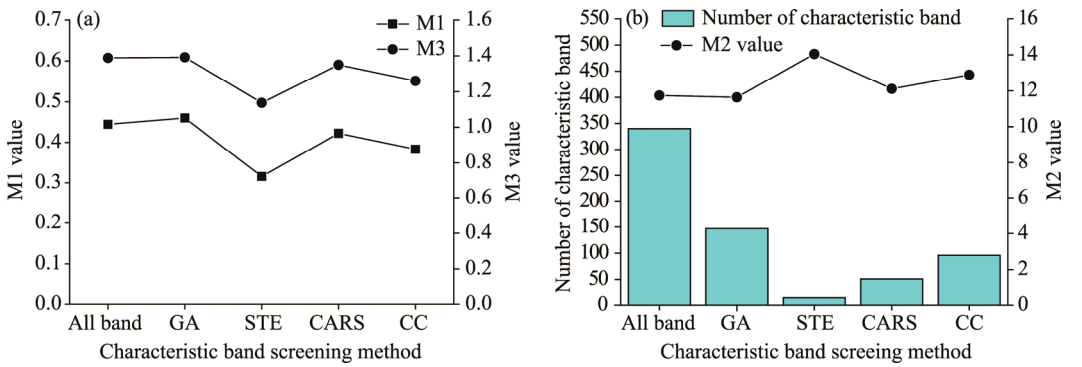

Fig. 7 M1 and M3 values of the different characteristic band screening methods (a), and M2 value of the different characteristic band screening methods and the numbers of characteristic bands in the models (b). M1, the mean value of the coefficient of determination (R^2) obtained by the leave-one-out cross-validation of 15 estimation models based on different transform spectra; M2, the mean value of the root-mean-squared error (RMSE) obtained by the leave-one-out cross-validation of 15 estimation models based on different transform spectra; M3, the mean value of the ratio of performance to deviation (RPD) obtained by the leave-one-out cross-validation of 15 estimation models based on different transform spectra; GA, genetic algorithm; STE, stepwise regression analysis; CARS, competitive adaptive reweighted sampling; CC, correlation coefficient method.

3.4 Optimal estimation model of the Au content in the soil

Table 2 indicates that the highest accuracy was obtained when the characteristic bands were selected from the transformed spectrum (1/ln$R$)' of *Seriphidium terrae-albae* using the CARS method and were input into the PLS method to construct an inversion model of the Au content in the soil. The R^2, RMSE, and RPD values obtained using leave-one-out cross-validation method constructed by 15 different transformed spectra for the model based on the CARS method were 0.8016, 7.07 μg/kg, and 2.20, respectively, and the number and distribution of characteristic bands included in the model are shown in Table 3. The number of characteristic bands was 56, accounting for only 16.50% of all bands in the range of 400–2400 nm. Compared with those of the model based on the full spectrum, the R^2 and RPD values of the model based on the characteristic band selected by the CARS method did not decrease significantly (0.8016 (model based on the CARS method) vs 0.8218 (model based on the full spectrum) for R^2, and 2.20 (model based on the CARS method) vs 2.37 (model based on the full spectrum) for RPD), and the RMSE did not significantly increase (7.07 vs 6.56 μg/kg); however, the number of bands used in the model was greatly reduced (56 for the model based on the CARS method *vs* 340 for the model based on the full spectrum), indicating that this

chinaXiv:202112.00013v1

ChinaXiv合作期刊

simplified model constructed based on the characteristic bands selected by the CARS method provided high accuracy, stability and generalization ability. This model is the best for the inversion of the Au content in the soil using the *Seriphidium terrae-albae* plantspectra.

Table 3 shows the distribution of characteristic bands in the spectral transformation $(1/\ln R)'$ screened with the CARS method. The selected characteristic bands were distributed in the visible-light, near-infrared, and short-wave infrared ranges. Among them, some of the bands that have been verified to be highly correlated with metal stress by other scholars (Hoque and Huntzler, 1992; Yang et al., 2016; Zhang et al., 2020) were screened and removed, including a band within the green peak range (542 nm), a band within the red valley range (652 nm), two bands within the red edge (692 and 772 nm), and a band within the water vapor absorption range (1457 nm). The above result suggests that the Au content estimation model constructed based on the characteristic bands selected by the CARS method has a strong physical significance.

Table 3 Distribution of characteristic bands in the spectral transformations (ST_{14} $(1/\ln R)'$) selected by the competitive adaptive reweighted sampling (CARS) method

Number of characteristic bands	Band distribution range (nm)
56	412–427, 517, 542, 577, 652, 692, 772, 842, 922, 947, 1032, 1047–1052, 1092, 1127, 1147, 1167–1172, 1197, 1212, 1402, 1427, 1457, 1487, 1622–1627, 1652, 1672, 1692–1697, 1702, 1712–1727, 1737, 1782, 1792, 2022, 2042, 2127, 2147–2152, 2167, 2197, 2242, 2257, 2292–2297, 2322–2327, 2342, 2352

4 Discussion

The emergence of hyperspectral technology is an important milestone in the development of remote sensing technology. Hyperspectral technology can be used to quickly detect tiny changes in plant spectra associated with metal stress, thus potentially providing quick and nondestructive estimates of the metal content. Currently, most scholars have used physical and statistical models to estimate the metal content (Liu et al., 2011, 2018; Hede et al., 2015; Zhang et al., 2017a, b, 2019; Zhang et al., 2018). Compared with physical models, statistical models have the advantages of being fast and simple; therefore, they have been widely used. However, although improved spectral resolutions can considerably enhance the amount of information and aid in the identification of characteristic parameters related to the degree of metal stress, the effects of interference factors such as soil background, canopy structure, and atmosphere make the canopy reflection spectra of plants extremely complex and variable (Cui et al., 2018). Additionally, statistical models could construct the quantitative relationship between the obtained plant spectrum (independent variable) and the metal content (dependent variable). A subtle error in the spectrum will often greatly affect the accuracy of the metal content estimates (Shi et al., 2016b; Wang et al., 2018). Therefore, eliminating background and noise effects while extracting useful spectral data has become a current research hotspot (Wang et al., 2018).

In this study, the Au content in the soil was estimated using plant spectra, and then 15 different forms of spectral transformation were explored to find the optimal spectral transformation form. The results indicate that the inversion model constructed using derivative spectroscopy provides a high accuracy. The reason for this phenomenon may be that the canopy reflectance spectrum of the plant canopy, rather than that of plant leaves, was measured in this study. Hyperspectral sensors are commonly mounted on aerospace and aviation platforms to vertically observe the ground and obtain the canopy spectra of plants. However, *Seriphidium terrae-albae* is a small shrub, and the leaves are small; thus, the canopy is not completely covered by leaves, which causes the canopy spectrum to contain noise, such as soil background information. In addition, canopy structure parameters, such as the leaf inclination of a plant, will affect the canopy reflectance spectrum. Due to the existence of this interference, some spectral information that contributes to the inversion of the Au content may be hidden. Thus, after the first-order and second-order derivatives are obtained, the soil background information and some random noise can be eliminated to some extent (Demetriades-Shah et al., 1990; Philpot, 1991; Smith et al., 2004). Moreover, the derivative spectra

chinaXiv:202112.00013v1

ChinaXiv合作期刊

can be used to extract the inflection point and slope information hidden in the plant spectra, amplify the differences between the spectra to a certain extent, and improve the inversion accuracy of the corresponding model (Cui et al., 2018).

It was found that the first derivative of reciprocal of logarithmic spectrum ((1/ln*R*)') is the best spectral transformation among the 15 spectral transform methods considered in this study. Notably, the plant canopy reflectance spectrum contains not only multiplicative noise but also additive noise. Calculating the reciprocal of the logarithm of the original spectrum can reduce the influence of multiplicative noise caused by the changes in light conditions to some extent (Bhargava and Mariam, 1992; Gong et al., 2001; Wang et al., 2011); afterwards, determining the first derivative can reduce the influence of additive noise (Zhang et al., 1997). Specifically, the spectral transformation (1/ln*R*)' not only reduces the effect of multiplicative noise but also decreases the effect of additive noise, allowing many tiny changes in the plant spectrum caused by the metal stress to be extracted. Consequently, among all the models considered, the metal content estimation model constructed with this transformed spectrum displays the highest accuracy.

5 Conclusions

In this study, the spatial coupling relationship between the Au content in the soil and concealed Au ore bodies was first analyzed, and the feasibility and effectiveness of estimating the Au content in the soil by using the spectra of plants in the visible, near-infrared, and short-wave infrared ranges were discussed. The results indicate that the Au content in the soil above an ore body is much higher than that in other areas, and the abnormal Au content in the soil corresponds well with the locations of ore bodies. Au content anomalies in the soil can be used as markers for the delineation of metallogenic targets. Then, the Au content in the soil can be accurately estimated by establishing a predictive model using the characteristic bands selected from plant spectra. However, in the modeling process, it is crucial to select the appropriate transform spectrum and band selection method. Among them, the PLS model based on the transformed spectrum (1/ln*R*)' and the CARS method perform well in estimating the Au content in the soil, with the R^2, RMSE, and RPD between the predicted value obtained by leave-one-out cross-validation method and measured value reaching 0.8016, 7.07 μg/kg, and 2.20, respectively. Overall, this study proposes a method to identify the concealed Au deposits by using plant spectra; this approach may have promising application prospects because of its simplicity and speed.

The greatest advantages of the proposed method are its simplicity and speed. Our research group has further constructed a "super low-altitude detection platform" that can obtain observations at the "submeter level" based on an airborne XT-912 power suspension glider and a HySpex hyperspectral sensor. Therefore, our next research goal is to apply the model constructed in this study in conjunction with the ultralow-altitude detection platform to determine whether the model is still valid. If feasible, the Au content in the soil could be quickly estimated over large areas, and metallogenic target areas could be remotely delineated. These advantages could greatly improve the efficiency of prospecting.

Acknowledgments

This research was funded by the National Natural Science Foundation of China (U1803117), the Young Scholars in Western China, Chinese Academy of Sciences (2020-XBQNXZ-014), the Tianchi Doctoral Plan (Y970000317), the Key Project of Natural Science Foundation of China-Xinjiang Joint Fund (U1803241), the Xinjiang Uygur Autonomous Region Talent Special Plan-Tianshan Outstanding Youth (2019Q033), and the Geological Exploration Project of Xinjiang Bureau of Geo-exploration and Minera development (XGMB202143).

References

Arias D. 1996. A case of successful soil geochemistry: the Rubiales Zn-Pb orebody (NW Spain). Journal of Geochemical Exploration, 56(3): 229–235.

chinaXiv:202112.00013v1

ChinaXiv合作期刊

Bandaru V, Daughtry, C S, Codling E E, et al. 2016. Evaluating leaf and canopy reflectance of stressed rice plants to monitor arsenic contamination. International Journal of Environmental Research and Public Health, 13(6): 606, doi: 10.3390/ijerph13060606.

Bhargava D S, Mariam D W. 1992. Cumulative effects of salinity and sediment concentration on reflectance measurements. International Journal of Remote Sensing, 13(11): 2151–2159.

Chakraborty S, Weindorf D C, Paul S, et al. 2015. Diffuse reflectance spectroscopy for monitoring lead in landfill agricultural soils of India. Geoderma Regional, 5: 77–85.

Cheng H, Shen R L, Chen Y Y, et al. 2019. Estimating heavy metal concentrations in suburban soils with reflectance spectroscopy. Geoderma, 336: 59–67.

Cui S C, Zhou K F, Ding R F, et al. 2018. Comparing the effects of different spectral transformations on the estimation of the copper content of *Seriphidium terrae-albae*. Journal of Applied Remote Sensing, 12(3): 036003, doi: 10.1117/1.JRS.12.036003.

Cui S C, Zhou K F, Ding R F, et al. 2021. Absorption and aggregation characteristics and changes in the reflectance spectrum of an arid desert plant under gold, copper, zinc and nickel stress. Natural Resources Research, 30(3): 2715–2731.

Demetriades-Shah T H, Steven M D, Clark J A. 1990. High resolution derivative spectra in remote sensing. Remote Sensing of Environment, 33(1): 55–64.

Dunagan S C, Gilmore M S, Varekamp J C. 2007. Effects of mercury on visible/near-infrared reflectance spectra of mustard spinach plants (*Brassica rapa* P.). Environmental Pollution, 148(1): 301–311.

Gong P, Pu R L, Yu B. 2001. Conifer species recognition: effects of data transformation. International Journal of Remote Sensing, 22(17): 3471–3481.

Han L, Chen R, Zhu H L, et al. 2020. Estimating soil arsenic content with visible and near-infrared hyperspectral reflectance. Sustainability, 12(4): 1476, doi: 10.3390/su12041476.

Hede A N H, Kashiwaya K, Koike K, et al. 2015. A new vegetation index for detecting vegetation anomalies due to mineral deposits with application to a tropical forest area. Remote Sensing of Environment, 171: 83–97.

Hong Y S, Chen Y Y, Yu L, et al. 2018. Combining fractional order derivative and spectral variable selection for organic matter estimation of homogeneous soil samples by VIS–NIR spectroscopy. Remote Sensing, 10(3): 479, doi: 10.3390/rs10030479.

Hoque E, Huntzler J S. 1992. Spectral blue-shift of red edge minitors damage class of beech trees. Remote Sensing of Environment, 39(1): 81–84.

Horler D N H, Barber J, Barringer A R. 1980. Effects of heavy metals on the absorbance and reflectance spectra of plants. International Journal of Remote Sensing, 1(2): 121–136.

Huang Z, Turner B J, Dury S J, et al. 2004. Estimating foliage nitrogen concentration from HYMAP data using continuum removal analysis. Remote Sensing of Environment, 93(1–2): 18–29.

Jarvis R M, Goodacre R. 2005. Genetic algorithm optimization for pre-processing and variable selection of spectroscopic data. Bioinformatics, 21(7): 860–868.

Jin J, Wang Q. 2019. Selection of informative spectral bands for PLS models to estimate foliar chlorophyll content using hyperspectral reflectance. IEEE Transactions on Geoscience and Remote Sensing, 57(5): 3064–3072.

Kemper T, Sommer S. 2002. Estimate of heavy metal contamination in soils after a mining accident using reflectance spectroscopy. Environmental Science & Technology, 36(12): 2742–2747.

Kooistra L, Wehrens R, Leuven R S E W, et al. 2001. Possibilities of visible–near-infrared spectroscopy for the assessment of soil contamination in river floodplains. Analytica Chimica Acta, 446(1–2): 97–105.

Kooistra L, Leuven R S E W, Wehrens R, et al. 2003. A comparison of methods to relate grass reflectance to soil metal contamination. International Journal of Remote Sensing, 24(24): 4995–5010.

Li H D, Liang Y Z, Xu Q S, et al. 2009. Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration. Analytica Chimica Acta, 648(1): 77–84.

Liu M L, Liu X N, Ding W C, et al. 2011. Monitoring stress levels on rice with heavy metal pollution from hyperspectral reflectance data using wavelet-fractal analysis. International Journal of Applied Earth Observation and Geoinformation, 13(2): 246–255.

Liu M L, Wang T J, Skidmore A K, et al. 2018. Heavy metal-induced stress in rice crops detected using multi-temporal Sentinel-2 satellite images. Science of the Total Environment, 637–638: 18–29.

Liu Z H, Lu Y, Peng Y P, et al. 2019. Estimation of soil heavy metal content using hyperspectral data. Remote Sensing, 11(12): 1464, doi: 10.3390/rs11121464.

Moros J, de Vallejuelo S F O, Gredilla A, et al. 2009. Use of reflectance infrared spectroscopy for monitoring the metal content of the estuarine sediments of the Nerbioi-Ibaizabal River (Metropolitan Bilbao, Bay of Biscay, Basque Country). Environmental Science & Technology, 43(24): 9314–9320.

chinaXiv:202112.00013v1

ChinaXiv合作期刊

Philpot W D. 1991. The derivative ratio algorithm: avoiding atmospheric effects in remote sensing. IEEE Transactions on Geoscience and Remote Sensing, 29(3): 350–357.

Ren H Y, Zhuang D F, Pan J J, et al. 2008. Hyper-spectral remote sensing to monitor vegetation stress. Journal of Soils and Sediments, 8(5): 323–326.

Sawut R, Kasim N, Abliz A, et al. 2018. Possibility of optimized indices for the assessment of heavy metal contents in soil around an open pit coal mine area. International Journal of Applied Earth Observation and Geoinformation, 73: 14–25.

Shi T Z, Chen Y Y, Liu Y L, et al. 2014. Visible and near-infrared reflectance spectroscopy—An alternative for monitoring soil contamination by heavy metals. Journal of Hazardous Materials, 265: 166–176.

Shi T Z, Liu H Z, Chen Y Y, et al. 2016a. Estimation of arsenic in agricultural soils using hyperspectral vegetation indices of rice. Journal of Hazardous Materials, 308: 243–252.

Shi T Z, Wang J J, Chen Y Y, et al. 2016b. Improving the prediction of arsenic contents in agricultural soils by combining the reflectance spectroscopy of soils and rice plants. International Journal of Applied Earth Observation and Geoinformation, 52: 95–103.

Shi T Z, Guo L, Chen Y Y, et al. 2018. Proximal and remote sensing techniques for mapping of soil contamination with heavy metals. Applied Spectroscopy Reviews, 53(10): 783–805.

Smee B W. 1998. A new theory to explain the formation of soil geochemical responses over deeply covered gold mineralization in arid environments. Journal of Geochemical Exploration, 61(1–3): 149–172.

Smith D B, Cannon W F, Woodruff L G. 2011. A national-scale geochemical and mineralogical survey of soils of the conterminous United States. Applied Geochemistry, 26: S250–S255.

Smith K L, Steven M D, Colls J J. 2004. Use of hyperspectral derivative ratios in the red-edge region to identify plant stress responses to gas leaks. Remote Sensing of Environment, 92(2): 207–217.

Song C A, Song W, Ding R F, et al. 2017. Phytogeochemical characteristics of *Seriphidium terrae-albae* (Krasch) Poljak in the metallic ore deposits in North part of East Junggar desert area, Xinjiang and their prospecting significance. Geotectonica et Metallogenia, 41(1): 122–132. (in Chinese)

Song W, Lei L Q, Song C A, et al. 2016. Characteristics of phytogeochemistry and prospecting choices of effective plants and elements in Kalatongke Cu-Ni ore field, Xinjiang. Journal of Guiling University of Technology, 36(2): 195–206. (in Chinese)

Sun W C, Zhang X. 2017. Estimating soil zinc concentrations using reflectance spectroscopy. International Journal of Applied Earth Observation and Geoinformation, 58: 126–133.

Sun W C, Zhang X, Sun X J, et al. 2018. Predicting nickel concentration in soil using reflectance spectroscopy associated with organic matter and clay minerals. Geoderma, 327: 25–35.

Tan K, Ma W B, Chen L H, et al, 2021. Estimating the distribution trend of soil heavy metals in mining area from HyMap airborne hyperspectral imagery based on ensemble learning. Journal of Hazardous, 401: 123288, 10.1016/j.jhazmat.2020.123288.

Timofeev I V, Kasimov N S, Kosheleva N E. 2016. Soil cover geochemistry of mining landscapes in the south-east of Transbaikalia (City of Zakamensk). Geography and Natural Resources, 37: 200–211.

Vega F A, Covelo E F, Andrade M L. 2006. Competitive sorption and desorption of heavy metals in mine soils: Influence of mine soil characteristics. Journal of Colloid and Interface Science, 298(2): 582–592.

Von Steiger B, Webster R, Schulin R. 1996. Mapping heavy metals in polluted soil by disjunctive kriging. Environmental Pollution, 94(2): 205–215.

Wang F H, Gao J, Zha Y, et al. 2018. Hyperspectral sensing of heavy metals in soil and vegetation: Feasibility and challenges. ISPRS Journal of Photogrammetry and Remote Sensing, 136: 73–84.

Wang J J, Wang T J, Shi T Z, et al. 2015. A wavelet-based area parameter for indirectly estimating copper concentration in *Carex* leaves from canopy reflectance. Remote Sensing, 7(11): 15340–15360.

Wang L, Bai Y L, Lu Y L, et al. 2011. Effect on retrieval precision for corn N content by spectrum data transformation. Remote Sensing Technology and Application, 26(2): 220–225. (in Chinese)

Wang R, Wu S, Wu K, et al. 2021. Estimation and spatial analysis of heavy metals in metal tailing pond based on improved PLS with multiple factors. IEEE Access, 9: 64880–64894.

Wang Z, Pan W, Yang X, et al. 2013. Soil geochemical anomaly characteristics and the geological significance of soil geochemical survey at Kavanga area in Kigoma, Tanzania. Contributions to Geology & Mineral Resources Research, 28(4): 634–640. (in Chinese)

Yang K M, Zhuo W, Zhang W W, et al. 2016. Study on the red edge response on derivative spectra of potted corn leaves stressed by lead ions. Science Technology and Engineering, 16(11): 110–114, 127. (in Chinese)

Yang X X, Luo X R, Zheng C J, et al. 2018. Geochemical characteristics of soil and prospecting direction in the Guoqing Area,

chinaXiv:202112.00013v1

ChinaXiv合作期刊

Northern Margin of the Hengyang Basin. Geology and Exploration, 54(4): 762–771. (in Chinese)

Yousefi G, Homaee M, Norouzi A A. 2018. Estimating soil heavy metals concentration at large scale using visible and near-infrared reflectance spectroscopy. Environmental Monitoring and Assessment, 190(9): 513, doi: 10.1007/s10661-018-6898-6.

Zhang C Y, Ren H Z, Qin Q M, et al. 2017a. A new narrow band vegetation index for characterizing the degree of vegetation stress due to copper: the copper stress vegetation index (CSVI). Remote Sensing Letters, 8(6): 576–585.

Zhang C Y, Ren H Z, Liang Y Z, et al. 2017b. Advancing the PROSPECT-5 model to simulate the spectral reflectance of copper-stressed leaves. Remote Sensing, 9(11): 1191, doi: 10.3390/rs9111191.

Zhang C Y, Ren H Z, Dai X J, et al. 2019. Spectral characteristics of copper-stressed vegetation leaves and further understanding of the copper stress vegetation index. International Journal of Remote Sensing, 40(12): 4473–4488.

Zhang C, Yang K M, Li Y, et al. 2020. Spectral characteristics and the study of pollution degree of maize leaves under copper and lead stress. Journal of the Indian Society of Remote Sensing, 48(1): 21–33.

Zhang L P, Zheng L F, Tong Q X. 1997. The estimation of vegetation variables based on high resolution spectra. Journal of Remote Sensing, 2(1): 111–114. (in Chinese)

Zhang Z J, Liu M L, Liu X N, et al. 2018. A new vegetation index based on multitemporal sentinel-2 images for discriminating heavy metal stress levels in rice. Sensors, 18(7): 2172, doi: 10.3390/s18072172.

chinaXiv:202112.00013v1